

When input fails to determine the grammar: Regional vs. Non-regional variation in Norwegian and Swedish

Abstract intended for European Dialect Syntax Workshop 5

This talk takes as its starting point in the results from the dialect survey taking place in Scandinavia at the moment in the ScanDiaSyn-project (see Lindstad et al. 2009). The goal of the talk is to raise the issue of two different types of variation that are apparent in the material: variation that is highly regionally conditioned (i.e. dialect variation) and variation that is not regionally conditioned, and from what we know, not sociologically conditioned either (i.e. idiolectal variation). I will give an overview of which linguistic phenomena fall into the two categories, and why. Secondly, I will present a case study of a phenomenon that falls into the second class: different types of long-distance anaphors in control infinitives and raise the question of how it is possible that two different grammars can arise in the same speech community.

When looking at variation within Mainland Scandinavia (here I will only discuss Norwegian and Swedish), the ScanDiaSyn data demonstrates that certain linguistic phenomena show a lot of geographically conditioned variation (for example word order in *wh*-questions), while variation for other types of phenomena (for example, the anaphoric system), is not geographically conditioned. In this talk, I will introduce measures for variation on the national level (N), the regional level (R = the county) and the local level (L = the individual measuring points) respectively. The measure shows how large a percentage of the informants give judgements that deviate from the most common judgement at the geographic level in question (i.e., N (national), R (region, county) and L (local)). A typical dialectal phenomenon should show noticeable variation on the national level, only very small variation on the regional level and almost no variation on the local level. Table 1 shows the values for two phenomena that show geographically conditioned variation: V3 in questions and *som*-insertion in subject *wh*-questions (the data below concern only Norwegian):

TEST SENTENCE	N%	R%	L%
Ka du heter? (“What you call?”)	51	18	11
Hvem som selger fiskeutyr her i bygda? (“Who that sells fishing equipment...”)	52	21	13

Table 2 shows the values for two phenomena that show variation that is only very mildly geographically conditioned: non-local binding of an object reflexive in an infinitival clause and binding of a reflexive possessive in the subject of an embedded clause by the subject of the matrix clause:

TEST SENTENCE	N%	R%	L%
Hun _i bad meg hjelpe seg _i (“She _i asked me to help REFL _i ”)	54	36	26
Regjeringen _i regner ikke med at forslaget sitt _i vil få flertall (“The g.ment don’t expect that SELF’S _i proposal...”)	52	42	30

In the talk I will try to find an explanation for the division of the data into the two types of variation (i.e. geographically and non-geographically conditioned variation). I will first rule out theoretically uninteresting explanations, like “poor informants” and “badly designed experiment”, and then connect to results from studies on language acquisition. In a study on the acquisition of anaphora in Danish, Jakubowicz (1992) shows that Danish children do not master long-distance anaphora until at least the age of 7.5 (while they show almost target like behavior when it comes to short distance anaphora already at the age of 3). Studies on the acquisition of word order in questions (see e.g. Westergaard 2009) have shown that children show target like behavior fairly early. I will hypothesize that there is a strong correlation between age of acquisition for a phenomenon, and the type of variation: a phenomenon that is acquired fairly late is not expected to show variation

that is highly geographically conditioned. This might simply be due to the obvious fact that the later you learn something, the more varied input you have got by the age of acquisition (i.e., input from books, TV and friends from far away). Alternatively, it might be the case that some rules are highly complex, and if the rules have a highly limited scope, and low usage frequency, the input might not be sufficient for fully determining the grammar. Rather, the chance or difference in learning strategies may influence the shape of the grammar.

The second part deals with variation in two types of long distance anaphora in Swedish: (1) binding of object reflexives in control infinitives, and (2) binding of possessive anaphors in the embedded object position in control infinitives:

- (1) Hon_i bad mig hjälpa sig_i.
 she asked me help RFLX
 ‘She asked me to help her.’
- (2) Hon_i bad mig passa sin katt_i
 she asked me to watch RFLX.POSS_i cat
 ‘She asked me to look after her cat’

The incoming data from the Swedish judgement tests shows that 69 per cent of the informants accept (2), while only 19 per cent accept (1) (mean value for (1): 3.9, mean value for (2): 2.3), and with some few exceptions, the speakers that accept (2) are a subset of the speakers that accept (1). This variation only is very mildly geographically conditioned. I will try to model the very subtle differences in the binding system between different speakers of Swedish, and also explain how these differences might arise. I will claim that speakers that accept both (1) and (2) treat the simple reflexive pronoun *sig* as a regular object pronoun, falling into the same paradigm as first and second person object pronouns *mig* and *dig*, and sharing the same syntactic properties with them. Speakers allowing only (2) on the other hand, must posit a special rule for the lexical item *sig*, stating that *sig* can only occur in the complement of a verb if it is co-referential with the external argument. I will argue that whether you end up with a grammar that includes (1) or not is probably determined by choice of learning strategy (i.e., if you are likely to stick to paradigms, or if you are likely to posit individual rules for particular lexical items) rather than input.

I hope that this talk will be a valuable contribution to the discussion of limits of dialectal variation, and I also hope that it can directly tie the field of micro-comparative syntax to the field of language acquisition in new and fruitful ways.